凸関数の平均について (作用素および作用素不等式の最近の話題)

藤井, 淳一

数理解析研究所講究録 1259 巻 2002 年 165-172

凸関数の平均について

大阪教育大学　藤井　淳一 (Jun Ichi Fujii)
Departments of Arts and Sciences (Information Science)
Osaka Kyoiku University

最近、Atteia-Raïssoul [3] によって (回帰的 Banach 空間上の) 凸汎関数の幾何平均が導入された。本質的には、まず調和平均を決めて、算術調和平均として定義されるものである cf. [6]。この手続きでは、帰納的に 2 つの単調増加・減少する汎関数列を定義して共通の極限として得られるものであるが、その収束条件については不明な点が多く、彼らが実際に扱っているのは、ごく限られた場合に過ぎないことが分かった。そこで、作用素平均の久保-安藤理論 [8] を利用して、Hilbert 空間 H 上の凸汎関数平均の一般論が出来ないかと考えた。一定度うまくいったと思われるので、ここに報告する。

まず、作用素平均を再確認しておこう。$B(H)$ の正作用素内の 2 項演算として、m が **作用素平均** であるとは、次の公理を満たすことである:

monotonicity: $A_1 \le A_2,\ B_1 \le B_2$ imply $A_1\ m\ B_1 \le A_2\ m\ B_2$.

semi-continuity: $A_n \downarrow A,\ B_n \downarrow B$ imply $A_n\ m\ B_n \downarrow A\ m\ B$.

transformer inequality: $T^*(A\ m\ B)T \le (T^*AT)m(T^*BT)$ (等号条件: $^\exists T^{-1}$).

normalization: $A\ m\ A = A$.

transformer inequality から、次はすぐに得られる：

homogeneity: $\alpha(A\ m\ B) = (\alpha A)m(\alpha B)$ for $\alpha > 0$.

この公理群のうち、凸関数自体は自然な積構造を持っていないねで、transformer inequality のみ定式化に注意が必要である。しかし、正数倍は自然に考えられるので、作用素平均と違って homogeneity との間にギャップが生じることが容易に想像される。この作用素平均については、

representing function $f_m(x) = 1\ m\ x$

によって、作用素単調関数との1対1対応がつき、その積分表示もえられる：

$$A\ m\ B = A^{1/2} f_m \left(A^{-1/2} B A^{-1/2}\right) A^{1/2} = aA + bB + \int_{(0,\infty)} (tA) : B \frac{1+t}{t} d\mu_m(t)$$

(ただし、μ は、$a = \mu_m(\{0\}), b = \mu_m(\{\infty\})$ となる確率測度) .
この積分表示の応用として、凸汎関数の平均を定義したい。

さて凸解析においては通常有限次元の実数空間が舞台になり、複素空間が扱われることはまれである。しかし複素 Hilbert 空間でも、内積の実数部分を取れば実 Hilbert 空間になり、理論上は何の障害もないし、無限次元の話もトポロジーを中心に論じられており、準備はある程度整っている。とはいえ、複素無限空間の凸解析についてはあまりまとまった記述もないので、この際きちんと準備をしよう。

有限次元の実空間でも、次のクラスは有用である [7, 11]: H 上の**下に有界な** (実数値) 凸汎関数 f の、 **(effective) domain** とは、

$$\mathrm{dom}\ f = \{\ x \in H \mid f(x) < \infty\ \}$$

のことであり、$\mathrm{dom}\ f = H$ のとき、**有限** と言われる。さらに、f が、**proper** とは、$\mathrm{dom}\ f \neq \emptyset$ のことを言い、**lower semi-continuous** もしくは、**closed** とは、次の **epigraph** が閉集合となっていることを言う:

$$\mathrm{epi}\ f = \{\ (x, \alpha) \in H \times \mathbb{R} \mid f(x) \leq \alpha\ \}$$

それで、有限次元同様、$\Gamma = \Gamma(H)$ を、下に有界で、proper な lower semi-continuous な H 上の凸汎関数全体の作るクラスとする。このクラスの汎関数は、いろんな意味でいい性質を共有しており、この中で平均を考察することにする。

このクラスに入る典型的な凸汎関数を挙げておこう。

Example.

1. **indicator** $\mathbf{1}_C(x) = \begin{cases} 0 & (x \in C) \\ \infty & \text{otherwise} \end{cases}$ $\quad C$: closed convex set

   当然ながら、これは有限ではない例である。

2. $f_A(x) = \dfrac{1}{2} \langle Ax, x \rangle \qquad A \in B(H), \quad A \geq 0.$

   これは、$f(\gamma x) = |\gamma|^2 f(x)$ という意味で、**quadratic** であり、有限である。

このクラス内の重要な演算として (複素空間なので、実部をとっているが)、**Legendre-Fenchel conjugate** f^* と呼ばれるものがある:

$$f^*(y^*) = \sup_x \mathrm{Re}\ \langle x, y\rangle - f(x)$$

これは、古くは Legendre transform と呼ばれているものであるが、Fenchel がその双対性についてめざましい結果を出したため、現在では上記のように呼ばれている。しかし、人によっては、Young-Fenchel conjugate と言うこともある。それはたとえば、

$$f(x) = |x|^p/p \quad ならば、 \qquad f^*(y^*) = |y|^q/q$$

となる事より、

Young's inequality: $f^*(y^*) \geq \mathrm{Re}\ \langle x, y\rangle - f(x)$

と呼ばれることに納得されるのではないだろうか。この双対凸汎関数の持つ性質は、

(1*) $f \leq g$ implies $f^* \geq g^*$.
(2*) $(f \pm \alpha)^* = f^* \mp \alpha$.
(3*) $(\alpha f)^*(\alpha y^*) = \alpha f^*(y^*) \quad (\alpha > 0)$.
(4*) f が、l.s.c でなくても f^* :l.s.c. (lower semi-continuous)
(5*) $f^{**} = f \in \Gamma$

である。lower semi-continuous でないような関数 f については、$f^{**} = \mathrm{cl}\ f$ となる。ここで、関数の閉包は、epi cl$f$ = cl (epi f) というように、epigraph の閉包をとった関数のことである。しかし、最も注目されるのは、上で例示した，正作用素から導かれる場合に

$$\exists A^{-1}のとき、 \qquad f_A^* = f_{A^{-1}}$$

となることを見れば，双対はある種の「逆元」であることが分かるだろう。

位相にも目を向けておこう。有限次元では，次の **Epi-convergence** E-lim とよばれる列収束がいい性質を持っている:

$$\mathrm{epi}\ (\operatorname*{E\text{-}lim}_{n\to\infty} f_n) = \lim_{n\to\infty} \mathrm{epi}\ f_n$$

$$f = \operatorname*{E\text{-}lim}_{n\to\infty} f_n \Longleftrightarrow f^* = \operatorname*{E\text{-}lim}_{n\to\infty} f_n^*$$

無限次元でも，これを拡張して，U.Mosco [9] は、次の収束概念を導入した:

Mosco 収束 $f = \underset{n\to\infty}{\text{M-lim}}\, f_n \quad \rightleftharpoons$

(i) $\quad \forall x \in H, \exists x_n \in H$ with $\underset{n\to\infty}{\text{s-lim}}\, x_n = x$ and $f(x) = \lim_{n\to\infty} f_n(x_n)$,

(ii) $\quad f(x) \le \liminf_{n\to\infty} f_n(x_n)$ for $\underset{n\to\infty}{\text{w-lim}}\, x_n = x$.

これも、次の性質を維持している:

$$f = \underset{n\to\infty}{\text{M-lim}}\, f_n \Longleftrightarrow f^* = \underset{n\to\infty}{\text{M-lim}}\, f_n^*$$

特に単調収束する場合，Mosco 収束することはすぐ分かる (cf. [4])。これらの事柄を背景にして，Atteia-Raïssouli は、調和平均を

$$f\tau_h g = \left(\frac{f^* + g^*}{2}\right)^*$$

と（明示的ではないが）定め，幾何平均 $f\tau g$ は、算術-調和平均として定めた [3]。同じ方法でいくつかの平均を作ることができる [10]。ただし定義があまり明確でなく，収束の関係上、dom f = dom g の場合に定義されているだけで、それ以外のどのような場合に定義できるかは不明である。そこで、一般の平均も定義したいので，まずは、**並列和** ([1, 5]) に対応する概念が凸解析で扱われていないかに注目した。すると、**inf-convolution** $f * g$

$$f * g(x) = \inf_{y+z=x} f(y) + g(z)$$

と呼ばれている概念があった（epi-sum と呼ばれることもある)。これは、並列和の次の公式

parallel sum $\qquad \langle A : Bx, x\rangle = \inf_{y+z=x} \langle Ay, y\rangle + \langle Bz, z\rangle$

を知っている人なら飛びつくだろう。しかも、正作用素に対応する凸汎関数の場合，

$$f_A * g_B = f_{A:B}$$

となることも

$$A : B = (A^{-1} + B^{-1})^{-1}, \quad \text{調和平均}\ AhB = 2(A : B)$$

に注意すれば，すぐ分かる。ところが、$2(f*f)\neq f$ であることは，

$$f*f(x)=2\inf_{y+z=x}\frac{f(y)+f(z)}{2}=2f\left(\frac{x}{2}\right)$$

から分かるし，さらに

$$\mathbf{1}_{C_1}*\mathbf{1}_{C_2}=\mathbf{1}_{C_1+C_2}\notin\Gamma$$

であるから、このクラス内の演算ですらないことまで分かった。そこで

$$\mathrm{cl}f*g\equiv(f*g)^{**}=(f^*+g^*)^*$$

となることは確認できるので，${}_Tf(x)=f(Tx)$ とするとき、$2(f:f)=f$ となるように係数を調整すれば，凸汎関数の**並列和**を

$$\begin{aligned}(f:g)(x)&=\frac{1}{4}\,{}_2((f^*+g^*)^*)(x)\equiv\frac{1}{4}(f^*+g^*)^*(2x).\\&=\frac{1}{4}(f*g)^{**}(2x)=\frac{1}{2}\left(\frac{f^*+g^*}{2}\right)^*(x)=\frac{1}{2}f\tau_h g.\end{aligned}$$

と定義すればよいが、結局 Atteia-Raïssouli 調和平均の半分以外に可能性がないという結論に達した。

作用素の場合と少し違うが、並列和は次のような「有界性」を持っているので、Γ 内では常に定義できるし、有限な凸汎関数の並列和も有限であることが分かる:

$t\in\mathrm{dom}\,f$, $s\in\mathrm{dom}\,g$ のとき

$$4(f:g)(x)\leq f(2x-s)+g(s),\quad\leq f(t)+g(2x-t).$$

さらに, $f(0)=0$ (resp., $g(0)=0$) なら

$$f:g(x)\leq\frac{1}{4}g(2x)\quad(\text{resp.,}\quad f:g(x)\leq\frac{1}{4}f(2x).)$$

とくに、f (resp., g): quadratic なら

$$f:g\leq g\qquad(\text{resp., } f:g\leq f)$$

さらに、この並列和も作用素平均同様、**monotonicity, semi-continuity** を満たす。半連続性については、Mosco 収束の性質が効いているのはいうまでもない。さらに、${}_Tf(x)=f(Tx)$ について、

$${}_Tf_A=f_{T^*AT}$$

がわかるので、transformer inequality が次のように定式化できる：

transformer inequality: ${}_T(f:g) \le ({}_Tf):({}_Tg)$.

homogeneity: $(\alpha f:\alpha g) = \alpha(f:g)$ for $\alpha > 0$.

quadratic preserving: If f and g is quadratic, then so is $f:g$.

ここで、T が可逆な場合の transformer equality は、$({}_Tf)^* = {}_{(T^{-1})^*}(f^*)$ から, homogeneity は (3*) の $(\alpha f)^*(\alpha y^*) = \alpha f^*(y^*)$ から示すことができる。この点、作用素平均の場合と多少関連が違っている。

以上のように、並列和を定義し、その性質は確保できたが、通常の「和」は定義の必要はないけれども、少し注意が必要である。以下、凸関数平均の一般論をする時には、

$$\mathrm{dom}\, f \cap \mathrm{dom}\, g \neq \emptyset$$

の仮定を置くことにする。なぜなら、この仮定を省くと「和」ですら proper でなくなり、Γ 内で定義できなくなるからである。久保-安藤理論の積分表示は、「平均とは、和と並列和の凸和」であることを示しているから、もはや、一般論の準備は整った。上記の仮定の下で、作用素平均 m に対する **凸関数平均** σ_m を

$$(f\sigma_m g)(x) = af(x) + bg(x) + \int_{(0,\infty)} \left({}_{\frac{1+t}{2}}f : t\,{}_{\frac{1+t}{2t}}g\right)(x)\frac{4}{1+t}d\mu_m(t)$$

とする。特に、f, g: quadratic のとき、

$$(f\sigma_m g)(x) = af(x) + bg(x) + \int_{(0,\infty)} ((tf):g)(x)\frac{1+t}{t}d\mu_m(t)$$

となるので、これは作用素平均の完全な拡張になっている。実際、${}_sf(x) = f(sx) = s^2f(x)$ $(s>0)$ より、

$$\left({}_{\frac{1+t}{2}}f : t\,{}_{\frac{1+t}{2t}}g\right)(x)\frac{4}{1+t} = \left(\frac{(1+t)^2}{4}f : \frac{(1+t)^2}{4t}g\right)(x)\frac{4}{1+t} = (tf:g)\frac{1+t}{t}$$

となる。この定義により、作用素平均同様、**monotonicity, semi-continuity, transformer inequality, homogeneity, quadratic preserving** を満たすことはすぐにわかる。この場合の半連続性は、もはや積分の単調収束定理である。平均の定義がややこしくなっているのは、normalization のためである。実際、

$$\begin{aligned}\frac{4}{1+t}\left(\tfrac{1+t}{2}f : t\tfrac{1+t}{2t}f\right)(x) &= \frac{1}{1+t}\left(\left(\tfrac{1+t}{2}f\right)^* + \left(t\tfrac{1+t}{2t}f\right)^*\right)^*(2x)\\ &= \frac{1}{1+t}\left(\left(\tfrac{1+t}{2}f\right)^* + t\left(\tfrac{1+t}{2}f\right)^*\right)^*(2x) = \frac{1}{1+t}\left((1+t)\left(\tfrac{1+t}{2}f\right)^*\right)^*(2x)\\ &= \left(\left(\tfrac{1+t}{2}f\right)^*\right)^*\left(\frac{2}{1+t}x\right) = \left(\tfrac{2}{1+t}f^*\right)^*\left(\frac{2}{1+t}x\right)\\ &= \tfrac{1+t}{2}(f^{**})\left(\frac{2}{1+t}x\right) = f^{**}(x) = f(x).\end{aligned}$$

によって、積分内が定数となり、測度が確率測度だから $f\sigma_m f = f$ が得られる。

最後に、作用素平均との対応を考察しておこう。そのためには、まず quadratic functional f がどんな場合に作用素から導かれたものになっているか確認しておく必要がある。たとえば、部分空間に対する indicator も quadratic であるが、有限ではないので作用素からは導き得ない。一般に、quadratic な凸汎関数 f は、$f(0) = 0$, $f(x) = f(-x)$ となり、

$$f(x) = \frac{f(x) + f(-x)}{2} \geq f(0) = 0$$

より、非負な値しかとらない。また、汎関数としての有界性を、仮に **locally bounded** と呼んでおくと:

$$\sup_{\|x\|=1} f(x) < \infty \iff f \text{ is continuous at } 0$$

が、線形な場合と同様に成り立ち、この場合当然ながら有限となる。さらに、作用している空間が内積空間となるためには、空間として中線定理が成り立つ必要があるので、汎関数としては、

$$f(x+y) + f(x-y) = 2\Big(f(x) + f(y)\Big)$$

という性質を持っている必要がある。このとき、f を、**PL functional** と呼ぶことにする。すると

$$4\Phi(x, y) = f(x+y) - f(x-y) + if(x+iy) - if(x-iy).$$

が連続準双線型形式となるので、$f = f_A$ となる $A \geq 0$ の存在がわかる。

作用素平均 m から導かれた凸汎関数平均 σ_m は、locally bounded, PL functional という性質は保存する。この性質を、**LBPL preserving** と呼んでおこう。すると、

Γ_0: locally. bdd. PL functionals in Γ

Σ: monotonicity, semi-continuity, homogeneity, transformer inequality,

normalization をみたす LBPL preserving な凸関数平均全体

としたとき、作用素平均と $\Sigma\big|_{\Gamma_0}$ との間に 1 : 1 対応がつくことがわかる。

また、Atteia-Raïssouli の平均が定義され、汎関数を quadratic なものに限るならば、ここで定義した平均と一致するが、幾何平均でさえも、適用される関数の範囲がこちらの場合のほうが広いことが、簡単な例から確かめられるので、一応当初の目的は達したことになる。

参考文献

[1] W.N.Anderson and R.J.Duffin: *Series and parallel addition of matrices*, J. Math. Anal. Appl., **26**(1969), 576–594.

[2] T.Ando: Topics on operator inequalities, Hokkaido Univ. Lecture Note, 1978.

[3] M.Atteia and M.Raïssouli: *Self dual operators on convex functionals geometric mean and square root of convex functionals*, J. Convex Anal., **8**(2001), 223–240.

[4] H.Attouch: Variational convergence for functions and operators, Pitman, 1984.

[5] P.A.Fillmore and J.P.Williams: *On operator ranges*, Adv. in Math., **7**(1971), 254–281.

[6] J.I.Fujii: *Arithmetico-geometric mean of operators*, Math. Japon., **23** (1978), 667–669.

[7] J-B.Hiriart-Urruty and C.Lemaréchal: Convex analysis and minimization algorithms I, II, Springer-Verlag, 1993.

[8] F.Kubo and T.Ando: *Means of positive linear operators*, Math. Ann., **248** (1980) 205–224.

[9] U.Mosco: *Convergence of convex sets and of solutions of variational inequalities*, Adv. in Math., **3**(1969), 510–585.

[10] M.Raïssouli and M.Chergui: *Arithmetico-Geometric and Geometrico-harmonic means of two convex functionals*, to appear in Sci. Math. Japon.

[11] R.T.Rockafellar and R. J-B. Wets: Variational analysis, Springer-Verlag, 1998.